NUMERICAL MATHEMATICS: Theory, Methods and Applications
Numer. Math. Theor. Meth. Appl., Vol. **1**, No. 1, pp. 78-91 (2008)

Some Remarks on the Convex Feasibility Problem and Best Approximation Problem

Qingzhi Yang[1,*] and Jinling Zhao[2]

[1] *School of Mathematics and LPMC, Nankai University, Tianjin 300071, China.*

[2] *School of Applied Science, University of Science and Technology, Beijing 100080, China.*

Received 11 February, 2007; Accepted (in revised version) 1 December, 2007

Abstract. In this paper we investigate several solution algorithms for the convex feasibility problem (CFP) and the best approximation problem (BAP) respectively. The algorithms analyzed are already known before, but by adequately reformulating the CFP or the BAP we naturally deduce the general projection method for the CFP from well-known steepest decent method for unconstrained optimization and we also give a natural strategy of updating weight parameters. In the linear case we show the connection of the two projection algorithms for the CFP and the BAP respectively. In addition, we establish the convergence of a method for the BAP under milder assumptions in the linear case. We also show by examples a Bauschke's conjecture is only partially correct.

AMS subject classifications: 49M37, 90C25, 90C90

Key words: Convex feasibility problem, best approximation problem, projection method, convergence.

1. Introduction

The convex feasibility problem (CFP) is to find a point in the nonempty intersection $C = \bigcap_{i=1}^{m} C_i$ of a family of closed convex subsets $C_i \subseteq R^n, 1 \le i \le m$, of the n-dimensional Euclidean space. It is a fundamental problem in many areas of mathematics and the physical sciences. More precisely, it has been used to model significant real-world problems including image reconstruction from projections, radiation therapy treatment planning, and crystallography (see [7] and the references therein). The convex sets $\{C_i\}_{i=1}^{m}$ represent mathematical constraints obtained from the modelling of the real-world problem.

The best approximation problem (BAP) is to find the projection of a given point $y \in R^n$ onto the nonempty intersection $C := \bigcap_{i=1}^{m} C_i \neq \emptyset$ of a family of closed convex subsets $C_i \subseteq R^n, 1 \le i \le m$, i.e., we need to look for a point in C which is closest to y. The relevant

*Corresponding author. *Email addresses:* `qz-yang@nankai.edu.cn` (Q. Yang), `twnnzhao@yahoo.com.cn` (J. Zhao)

background knowledge may consult [1] and [16]. In the CFP, any point in the intersection is acceptable to the real-world, while for the BAP it is appropriate if some point $y \in R^n$ has been obtained from modelling and computational efforts that initially did not take into account the constraints represented by the sets $\{C_i\}_{i=1}^m$ and now one wishes to incorporate them by seeking a point in the intersection of the convex sets which is closest to the point y.

For the CFP a number of solution methods have been presented (see [3, 8–11, 13, 14, 21–23, 27, 28, 30]). Among them, some are particularly designed for the CFP of special forms. Roughly speaking, these algorithms can be divided into two categories: projection method and interior method. For the BAP, several projection-type algorithms have been proposed to solve it. (see [2, 7, 16–18]).

The orthogonal projection $P_\Omega(z)$ of a point $z \in R^n$ onto a closed convex set $\Omega \subseteq R^n$ is a point of Ω defined by

$$P_\Omega(z) := arg\min\{\|z - x\|_2\},$$

where $\|.\|_2$ is the Euclidean norm in R^n.

The projection-type methods employ projection onto the individual convex sets in order to reach the required point in the intersection. Obviously the solution of the BAP for any given y is a solution of the CFP provided that $\bigcap_{i=1}^m C_i \neq \emptyset$. So it is easy to see that the iterate projection algorithms for the BAP are usually more complicated than algorithms for the CFP. However, we will show in Section 3 that at least in the linear case a relaxed projection algorithm for the CFP will produce a solution of the BAP as long as taking y as the starting point of the iteration.

In the present paper we intend to supply a relatively unified treatment for various projection algorithms for the CFP based on the steepest descent method. We also study the iterate behaviors of the sequential and simultaneous versions of Halpern-Lions-Wittmann-Bauschke (HLWB) algorithm for the BAP. In particular, we establish the convergence of the simultaneous HLWB in the linear case, which means that the algorithms can be accelerated. Moreover we show that when $\bigcap_{i=1}^m C_i = \emptyset$ a Bauschke's conjecture is only partially correct.

This paper is organized as follows. In Section 2, based on the reformulations of the CFP we naturally deduce the exact and surrogate relaxed projection algorithms for the CFP, from which we suggest a more natural updating strategy of weight parameters. In Section 3, we prove the convergence of the simultaneous HLWB algorithm in the linear case under mild conditions. In Section 4, we discuss the simultaneous HLWB in the case of intersection sets being empty and show that a conjecture due to Bauschke is only partially correct.

2. Several algorithms for the CFP

In this section three well-known algorithms for the CFP are further discussed.

2.1. The projection algorithms for the CFP revisited

It is known that for a closed convex set in R^n, there is a unique convex function associated to this set (see, e.g., [28]). Denote

$$C_i = \{x \in R^n : f_i(x) \leq 0\}, \quad i = 1, \cdots, m,$$

where f_i, $i = 1, \cdots, m$ are convex. Our motivation here is that using the steepest descent method, but based on the different reformulations of the CFP, yields two known algorithm. Particularly, we give a natural scheme of updating weight parameters for the surrogate project algorithm for the CFP.

First we recall the steepest descent method for solving the unconstrained optimization problem: min $f(x)$, where $f(x)$ is a continuously differential function in R^n. Then the steepest descent method can be stated as follows: Given arbitrary x^o, for $k = 0, 1, \cdots$, calculate

$$x^{k+1} = x^k - t_k \nabla f(x^k),$$

where t_k is a suitable stepsize. (When $f(x)$ is convex, its subgradient, denoted by $\partial f(x)$, exists. At this time, the steepest descent method solving min $f(x)$ is to replace the $\nabla f(x^k)$ by an element ξ^k of $\partial f(x^k)$.)

We know that when $f(x)$ is differentiable, if t_k is an exactly optimal stepsize for each k, then the sequence generated by this algorithm converges to a stationary point provided the level set of $f(x)$ is bounded, i.e., for every α, $\{x : f(x) \leq \alpha\}$ is bounded. However, in the situation of $f(x)$ being nondifferentiable, the limit point of the sequence generated by the steepest descent method may not be a optimal solution even if it exists (see, e.g., [5]). Since the exact stepsize is not realistic in general, various inexact linesearch strategies have be proposed to guarantee the convergence of corresponding algorithms. Because the reformulated problems of the CFP possess specific properties generally, such as the zero value of optimal objective, it is more flexible to determine the inexact stepsize to ensure the convergence of the algorithm.

Let ω be given such that

$$\omega_i > 0, \;\; i = 1, \cdots, m, \quad \sum_{i=1}^{m} \omega_i = 1.$$

Then we have the following reformulation of the CFP:

$$(R_1) \qquad \min \tfrac{1}{2} \sum_{i=1}^{m} \omega_i \|x - P_{C_i}(x)\|_2^2.$$

Since it is known (see, e.g., [6]) that

$$\nabla\Big(\frac{1}{2}\sum_{i=1}^{m} \omega_i \|x - P_{C_i}(x)\|_2^2\Big) = \sum_{i=1}^{m} \omega_i (x - P_{C_i}(x)) = x - \sum_{i=1}^{m} \omega_i P_{C_i}(x),$$

applying steepest descent method to (R_1) gives the following iteration for solving the CFP:

$$x^{k+1} = x^k - t_k\Big(x^k - \sum_{i=1}^{m} \omega_i P_{C_i}(x^k)\Big) = (1 - t_k)x^k + t_k \sum_{i=1}^{m} \omega_i P_{C_i}(x^k). \tag{2.1}$$

Setting $\theta_k = 1 - t_k$ yields the so-called relaxed combination projection algorithm. Of course, due to the difficulty of calculating the exact projection, the most popular idea in the projection-type algorithms is to replace the exact projection in the above scheme by some surrogate projection such as the half-space projection (see [27, 29] and the relevant references therein). To see how to derive the relaxed combination projection algorithm with the half-space surrogate projection naturally, let us consider the following reformulation of the CFP:

$$(R_2) \qquad \min \ \tfrac{1}{2}\sum_{i=1}^{m} \omega_i \max{}^2(f_i(x), 0).$$

Obviously the $\bigcap_{i=1}^m C_i \neq \emptyset$ if and only if (R_2) has the optimal objective value of zero. It has been known that

$$\partial\Big(\frac{1}{2}\sum_{i=1}^{m}\omega_i \max{}^2(f_i(x),0)\Big) = \sum_{i=1}^{m}\omega_i \max(f_i(x),0)\partial f_i(x),$$

where $\partial f_i(x)$ denotes the subgradient of f_i at x, i.e.,

$$\partial f_i(x) = \{\eta : f(y) \geq f(x) + \langle \eta, y - x\rangle, \forall y\}, \quad 1 \leq i \leq m.$$

Let ξ_i stand for an element in $\partial f_i(x)$. If we let ω be varied from the current iteration to the next iteration and assume that at kth iteration, $\omega := \omega^k$, then applying the steepest descent method to (R_2) leads to the following scheme:

$$x^{k+1} = x^k - t_k \sum_{i=1}^{m} \omega_i^k \max(f_i(x^k), 0)\xi_i^k, \tag{2.2}$$

where t_k is suitable stepsize and $\xi_i^k \in \partial_i(x^k)$, for $i = 1, \cdots, m$. From this scheme and taking into account the optimal condition of the solution, we see that

$$\omega_i^{k+1} = \frac{\omega_i^k \max(f_(x^k), 0)}{\sum_{j=1}^{m} \omega_j^k \max(f_j(x^k), 0)}, \quad i = 1, \cdots, m$$

is a natural strategy of updating ω^k provided that

$$\sum_{j=1}^{m} \omega_j^k \max(f_j(x^k), 0) \neq 0.$$

For each $i = 1, \cdots, m$, define the half space associated with $f_i(x)$ at x^k

$$H_i^k = \{x : f_i(x^k) + \langle \xi_i^k, x - x^k\rangle \leq 0\}.$$

We know that the projection of x^k onto H_i^k is

$$P_{H_i^k}(x^k) = x^k - \frac{\max(f_i(x^k), 0)}{\|\xi_i^k\|_2^2}\xi_i^k.$$

Thus (2.2) can be rewritten as

$$x^{k+1} = x^k + t_k \sum_{i=1}^{m} \omega_i^k \|\xi_i^k\|_2^2 (P_{H_i^k}(x^k) - x^k). \tag{2.3}$$

If we denote

$$\theta_i^k = \frac{\omega_i^k \|\xi_i^k\|_2^2}{\sum\limits_{j=1}^{m} \omega_j^k \|\xi_j^k\|_2^2}, \quad i = 1, \cdots, m, \tag{2.4}$$

and

$$\rho_k = t_k \sum_{j=1}^{m} \omega_j^k \|\xi_j^k\|_2^2.$$

Then we obtain from (2.3) the following iteration

$$x^{k+1} = x^k + \rho_k \sum_{i=1}^{m} \theta_i^k (P_{H_i^k}(x^k) - x^k). \tag{2.5}$$

This is actually the common form of the relaxed projection algorithm for solving the CFP, see, e.g., [27]. The above argument provides a new insight of the relaxed projection algorithm. Specially, it is not easy to get a natural scheme of updating weight parameter θ^k from only (2.5). Several strategies for updating θ^k is seemly rather artificial and not closely related to iterate step, see [3, 27, 29] and the references therein. In fact, from the above argument we see that the natural way of updating θ^k should be (2.4). Of course, when $\theta_i^k = 0$, i.e., $\omega_i^k = 0$, a restarting strategy for θ_i is necessary, at least for the convergence proof. When this situation occurs, it is suggested to use the way as used in [27]. The above argument also tells us that slow convergence rate may occur for (2.5), since it is actually an application of the steepest descent method. This motivates us to study fast algorithms for the CFP when possible.

2.2. Conjugate Gradient method for the CFP

To speed up the convergence rate of the algorithms discussed in the previous subsection, we intend to show that two fast methods for solving unconstrained optimization: Conjugate Gradient (CG) method and nonsmooth Newton method which may be used to solve the CFP under certain conditions.

In [38], we show that the following CG method can be applied to the CFP provided that the projections in reformulation (R_1) of the CFP are easy to calculate.

Algorithm 2.1. Let x^0 be an arbitrary vector in R^n. For $k = 0, 1, \cdots$, calculate

$$x^{k+1} = x^k + \alpha_k d^k, \tag{2.6}$$

where

$$d^k = \begin{cases} -g^k & k = 0, \\ -g^k + \beta_k d^{k-1} & k > 0, \end{cases} \tag{2.7}$$

$$g^k = 2\sum_{i=1}^{m} \omega_i (x^k - P_{C_i}(x^k)). \tag{2.8}$$

The various choices of β_k lead to different conjugate gradient (CG) methods, see, e.g., [38]. For examples, if

$$\beta_k = \beta_k^{FR} = \frac{\|g^k\|_2^2}{\|g^{k-1}\|_2^2},$$

we obtain F-R (Fletcher-Reeves) conjugate gradient method; if

$$\beta_k = \beta_k^{PRP} = \frac{(g^k)^T (g^k - g^{k-1})}{\|g^{k-1}\|_2^2},$$

we obtain P-R-P (Polak-Ribière-Polyak) conjugate gradient method. As to α_k, we can use exact or inexact linear search to obtain it. In fact we have the following (see [38])

Theorem 2.1. *Assume that the objective function $f(x)$ is bounded below, and $\nabla f(x)$ is Lipschitz continuous with modulus $L > 0$. If $0 < \alpha_k \equiv \eta < \frac{1}{4L}$ in P-R-P method, then it holds that $\nabla f(x^k)^T d^k < 0$ for all k. In this case, the algorithm defined by* (2.6)-(2.8) *converges globally in the sense that*

$$\lim_{k \to +\infty} \|\nabla f(x^k)\|_2 = 0.$$

It is easy to see the requirements in the above theorem are satisfied by (R_1). In [38], our preliminary experiments show that the CG method is superior to the CQ algorithm for some randomly selected split feasibility problem (SFP). It is well-known that the SFP is a particular case of the CFP (see [6] and the references therein) and the CQ algorithm is an application of the steepest descent method with fixed stepsize (see [38]). Likewise, it seems promising to use Algorithm 2.1 to solve the CFP based on (R_1) as long as the associated projections can be easily obtained.

3. Simultaneous HLWB for the BAP in the linear case

In this section, we are concerned with the BAP. Given a vector y in R^n, the BAP is to find a point in intersection sets of some closed convex sets of R^n, which is closest to y. Denote $C = \bigcap_{i=1}^m C_i$. The BAP is to solve the following specific optimization problem:

$$(P_1) \qquad \min \tfrac{1}{2}\|x-y\|_2^2 \quad s.t. \;\; x \in \bigcap_{i=1}^m C_i.$$

As pointed out in [7], the iterative projection algorithms for the BAP, such as the algorithms of Dykstra and others, are more complicated than algorithms for the CFP because they must have, in their iterative steps, some built-in "memory" mechanism to the original point whose projection is sought after. However, at least for the linear case, we show below that the iterative scheme (2.1) for the CFP can be applied to the BAP with y as the starting point of the procedure.

Let $\omega > 0$ be the weight parameter. Then the simultaneous HLWB for the BAP is as follows (see, e.g., [2, 7, 12])

$$x^{k+1} = \rho_k y + (1-\rho_k)\sum_{i=1}^m \omega_i P_{C_i}(x^k). \tag{3.1}$$

Obviously, this iteration is easily executed. In [7], Censor carried out several experiments based on several algorithms for the BAP, including the sequential and simultaneous HLWB, without theoretical analysis. From Censor's experiments we see that the simultaneous HLWB has much slow convergence rate compared with the sequential HLWB. It is quite strange, since the simultaneous algorithm seems a good way to speed up the convergence of the sequential algorithm. There are two possible factors responsible for the slow convergence: one is the slow rate of ρ_k approaching to 0, and another one is the constant weight parameter. Probably using the dynamic strategies of updating ω, such as that in [27] or that mentioned in the previous section for the CFP, is preferred. Of course, this may lead to difficulty in analyzing the convergence of the algorithm. Below, we will establish the convergence of the iteration (3.1) in the linear case with less requirements.

In [2, 7], $\{\rho_k\}$ is required to satisfy following conditions:

(C1) $\rho_k \to 0$ as $k \to +\infty$;

(C2) $\sum_{k=0}^{\infty} \rho_k = +\infty$;

(C3) For any given positive integer m, $\sum_{k=0}^{\infty} |\rho_k - \rho_{k+m}| < +\infty$.

In order to speed up the convergence rate of the algorithm, it will be significant to remove the conditions (C2) and (C3). For the convenience of the analysis, first we transform the BAP into a simpler equivalent form.

Denote $z = x - y$. Then (P_1) is equivalent to finding the projection of the origin onto $\bigcap_{i=1}^m \hat{C}_i$, where $\hat{C}_i = C_i - \{y\}$ for $i = 1, \cdots, m$, i.e.,

$$(P_2) \qquad \min \|z\|_2^2 \quad s.t. \;\; z \in \bigcap_{i=1}^m \hat{C}_i.$$

Theorem 3.1. *Let $\hat{C}_i = \{z \in R^n : a_i^T z = b_i\}$, where a_i is a family of standardly orthogonal vectors, i.e., $\|a_i\|_2 = 1,\ i = 1, \cdots, m$ and $a_i^T a_j = 0,\ i \neq j$. If $\{\rho_k\}$ satisfy* (C1)*, then the sequence generated by* (3.1) *converges to the solution of the BAP.*

Proof. Obviously the assumption implies that $m \le n$, because for each iterate point z^k,

$$P_{\hat{C}_i}(z^k) = z^k - (a_i^T z^k - b_i)a_i.$$

Then at this time the simultaneous HLWB for solving (P_2) can be written as follows:

$$z^{k+1} = (1-\rho_k)\sum_{i=1}^{m} \omega_i \Big(z^k - (a_i^T z^k - b_i)a_i\Big).$$

Denote $A = (a_1, \cdots, a_m)^T$, $b = (b_1, \cdots, b_m)^T$, $A(\omega) = (\sqrt{\omega_1}a_1, \cdots, \sqrt{\omega_m}a_m)^T$ and $b(\omega) = (\sqrt{\omega_1}b_1, \cdots, \sqrt{\omega_m}b_m)^T$. Hence the above iteration scheme can be written as

$$z^{k+1} = (1-\rho_k)\Big(z^k - A^T(\omega)A(\omega)z^k + A^T(\omega)b(\omega)\Big)$$

with the initial point $z^0 = 0$. Because

$$A(\omega)A^T(\omega) = diag(\omega_1, \cdots, \omega_m),$$

we know that from matrix theory that there exists an orthogonal matrix G such that

$$G^T A^T(\omega)A(\omega)G = (A(\omega)G)^T(A(\omega)G) = diag(\omega_1, \cdots, \omega_m, \overbrace{0, \cdots, 0}).$$

It directly follows that $e_j^T(A(\omega)G) = 0$ for $j = m+1, \cdots, n$. Therefore we obtain that

$$G^T z^{k+1} = (1-\rho_k)\Big((I - G^T A^T(\omega)A(\omega)G)(G^T z^k) + G^T A^T(\omega)b(\omega)\Big).$$

More precisely,

$$e_i^T G^T z^{k+1} = (1-\rho_k)\Big((1-\omega_i)e_i^T(G^T z^k) + e_i^T(A(\omega)G)b(\omega)\Big), \quad i = 1, \cdots, m$$

and

$$e_j^T G^T z^{k+1} = (1-\rho_k)e_j^T(G^T z^k), \quad j = m+1, \cdots, n.$$

By recursion and taking into account $z^0 = 0$, we have for $i = 1, \cdots, m$,

$$e_i^T G^T z^{k+1} = s_i^k e_i^T (A(\omega)G)^T b(\omega)), \tag{3.2}$$

where

$$\begin{aligned} s_i^k &= (1-\rho_k) + (1-\omega_i)(1-\rho_k)(1-\rho_{k-1}) \\ &\quad + \cdots + (1-\omega_i)^k(1-\rho_k)(1-\rho_{k-1})\cdots(1-\rho_0). \end{aligned} \tag{3.3}$$

Since $\rho_k \to 0+$ and $0 < 1-\omega_i < 1$, the infinite series (3.3) is convergent, so is the sequence $\{e_i^T(G^T z^k)\}$ for $i = 1, \cdots, m$.

For $j = m+1, \cdots, n$, it is easy to see that $e_j^T(G^T z^k) \equiv 0$. It can be verified that

$$s_i^k = (1-\rho_k)(1 + (1-\omega_i)s_i^{k-1}).$$

Let $\hat{s}_i$ be the limit point of $\{s_i^k\}$. Thus one gets

$$\hat{s}_i = 1 + (1-\omega_i)\hat{s}_i.$$

Therefore $\hat{s_i} = 1/\omega_i$. It immediately follows from (3.2) that

$$e_i^T G^T z^{k+1} = s_i^k e_i^T (A(\omega)G)^T b(\omega)) \to e_i^T (A(\omega)G)^T b(\omega))/\omega_i \quad \text{as} \quad k \to \infty.$$

Denote the accumulation point of z^k by z^*. Then we conclude that $z^* = A^T b$. By the KKT conditions for the constrained optimization, it is easy to see that $z^* = A^T b$ is the optimal solution of the BAP

$$\min \ \frac{1}{2}\|z\|_2^2 \qquad s.t. \ a_i^T z = b_i, \ \ i = 1, \cdots, m.$$

This completes the proof of the theorem. $\square$

Corollary 3.1. *In the above theorem, if we denote*

$$\hat{C}_i = \{z \in R^n : a_i^T z \le b_i\}, \quad i = 1, \cdots, m$$

and assume that the sequence $\{z^k\}$ never falls into the interior of $\hat{C}_i$ for each $i = 1, \cdots, m$, then the sequence $\{z^k\}$ converges to the solution of the BAP (P_2).

Remark 3.1. *In fact, as long as $a_i (i = 1, \cdots, m)$ are linearly independent, Theorem* 3.1 *and its corollary still hold. Therefore we may let $\rho_k \equiv 0$ in order to speed up the convergence of the algorithm in this case. It is expected that Theorem* 3.1 *holds as long as each $\hat{C}_i$ is a polyhedron.*

In the following, we show that the relaxed combination projection algorithm (2.1) for solving the CFP can be applied to the BAP under certain conditions. Let us rewrite (2.1) as

Algorithm 3.1. Given a positive sequence $\{\rho_k\}$ such that $\rho_k \in (0, 1)$ and an arbitrary initial point x^0, for $k = 1, 2, \cdots$, calculate

$$x^{k+1} = \rho_k x^k + (1 - \rho_k) \sum_{i=1}^{m} \omega_i P_{C_i}(x^k).$$

It has been known that the sequence generated by Algorithm 3.1 converges to a point in $\bigcap_{i=1}^m C_i$ under certain conditions provided $\bigcap_{i=1}^m C_i \neq \emptyset$, see, e.g., [7]. It is natural to ask whether or not the limit point of the sequence generated by Algorithm 3.1 is the projection of the starting point onto the intersection. If it is the case, the rich results for the CFP may be helpful for the research of the BAP. In fact we have the following

Theorem 3.2. *Assume $\hat{C}_i = \{z \in R^n : a_i^T z = b_i\}$, a_i, $i = 1, \cdots, m$, is linearly independent vectors. If $\{\rho_k\}$ satisfy condition* (C1)*, then the sequence $\{x^k\}$ generated by Algorithm* 3.1 *with the origin as the starting point converges to the solution of (P_2).*

The proof of this theorem is similar to that of Theorem 3.1 and we omit it. Furthermore, we have the following conjecture.

Conjecture 3.1. *Suppose that $\{\rho_k\}$ satisfy condition* (C1) *and $\bigcap_{i=1}^m C_i \neq \emptyset$. Set y as the initial point of the scheme, then the sequence produced by Algorithm* 3.1 *converges to the projection of y onto $\bigcap_{i=1}^m C_i$.*

4. Some results when $\bigcap_{i=1}^m C_i = \emptyset$

Throughout this section we assume $\bigcap_{i=1}^m C_i = \emptyset$, i.e., the CFP is an inconsistent system. It is already well-known that when $\bigcap_{i=1}^m C_i \neq \emptyset$, both the sequential and simultaneous HLWB are convergent as long as the positive sequence $\{\rho_k\}$ satisfies conditions (C1)-(C3) in Section 3.

However, in practical applications, it is unknown in advance if $\bigcap_{i=1}^m C_i$ is empty. It is interesting to know the properties of the sequential and simultaneous HLWB or Algorithm 3.1 when $\bigcap_{i=1}^m C_i = \emptyset$.

In [25], Iusem and De Pierro gave several conditions which ensure that $\{x : x = \sum_{i=1}^m \omega_i P_{C_i}(x)\}$ is nonempty.

Theorem 4.1. *Assume* $\bigcap_{i=1}^m C_i = \emptyset$ *and*

$$\left\{x : x = \sum_{i=1}^m \omega_i P_{C_i}(x)\right\} \neq \emptyset.$$

If $\rho_k \in (0,1)$ *satisfies*

$$\sum_{k=1}^{\infty}(1-{\rho_k}^2) = +\infty, \tag{4.1}$$

then the sequence generated by Algorithm 3.1 *converges to a fixed point of the operator* $\sum_{i=1}^m \omega_i P_{C_i}$.

Proof. Let x^* be a fixed point of $\sum_{i=1}^m \omega_i P_{C_i}$, i.e.,

$$x^* = \sum_{i=1}^m \omega_i P_{C_i}(x^*).$$

Denote $N_i = I - P_{C_i}$, $i = 1, \cdots, m$, where I is the identity operator. It is already known that

$$\langle N_i(x) - N_i(y), x - y\rangle \geq \|N_i(x) - N_i(y)\|_2^2 \tag{4.2}$$

for any $x, y \in R^n$. Then we have that

$$\begin{aligned}
&\|x^{k+1} - x^*\|_2^2 \\
=\ & \|(x^k - x^*) - (1-\rho_k)(x^k - \sum_{i=1}^m \omega_i P_{C_i}(x^k))\|_2^2 \\
=\ & \|(x^k - x^*) - (1-\rho_k)\sum_{i=1}^m \omega_i N_i(x^k)\|_2^2 \\
=\ & \|x^k - x^*\|_2^2 - 2(1-\rho_k)\left\langle x^k - x^*, \sum_{i=1}^m \omega_i(N_i(x^k) - N_i(x^*))\right\rangle + S_k
\end{aligned}$$

$$\begin{aligned}
&\leq \|x^k - x^*\|_2^2 - 2(1-\rho_k)\sum_{i=1}^m \omega_i \|N_i(x^k) - N_i(x^*)\|_2^2 + S_k \\
&\leq \|x^k - x^*\|_2^2 - 2(1-\rho_k)\|\sum_{i=1}^m \omega_i (N_i(x^k) - N_i(x^*))\|_2^2 + S_k \\
&= \|x^k - x^*\|_2^2 - (1-{\rho_k}^2)\|\sum_{i=1}^m \omega_i (N_i(x^k) - N_i(x^*))\|_2^2,
\end{aligned} \tag{4.3}$$

where

$$S_k = (1-\rho_k)^2 \|\sum_{i=1}^m \omega_i (N_i(x^k) - N_i(x^*))\|_2^2.$$

Since $\rho_k \in (0,1)$, it follows that $\{\|x^k - x^*\|_2^2\}$ is monotonically decreasing and therefore is convergent and bounded. Moreover, it follows from (4.1) that

$$\inf \|\sum_{i=1}^m \omega_i (N_i(x^k) - N_i(x^*))\|_2^2 = \inf \|\sum_{i=1}^m \omega_i N_i(x^k)\|_2^2 = 0. \tag{4.4}$$

Assume that $\hat{x}$ is an accumulation point of $\{x^k\}$ such that $\sum_{i=1}^m \omega_i N_i(\hat{x}) = 0$, i.e., $\hat{x} = \sum_{i=1}^m \omega_i P_{C_i}(\hat{x})$. Thus $\hat{x}$ is a fixed point of the operator $\sum_{i=1}^m \omega_i P_{C_i}$ and we put it in place of x^* in the previous argument. Since $\{\|x^k - \hat{x}\|_2^2\}$ is convergent and a subsequence of it converges to zero, we have that the entire sequence converges to $\hat{x}$. $\square$

For the sequential HLWB, Bauschke [2] made a conjecture that $\|x^k\|_2 \to \infty$ as $k \to \infty$ when $\bigcap_{i=1}^m C_i = \emptyset$. The following examples show that this conjecture is only partially true.

Example 4.1. Let $C_1 = \{(x_1, x_2) : x_1 \geq 1\}$, and $C_2 = \{(x_1, x_2) : (x_1+1)^2 + x_2^2 \leq 1\}$. Obviously, $C_1 \bigcap C_2 = \emptyset$. Set the initial point $x^0 = (\frac{1}{3}, 1)$. One easily verifies that the sequence $\{x^k\}$ generated by the sequential HLWB has two accumulation points $\hat{x}^1 = (1,0)$ and $\hat{x}^2 = (0,0)$. More precisely, $x^{2l-1} \to \hat{x}^1$ and $x^{2l} \to \hat{x}^2$ as $l \to \infty$.

Example 4.2. Let $C_1 = \{(x_1, x_2) \geq (0,0) : x_1 x_2 \geq 1\}$ and $C_2 = \{(x_1, x_2) : x_2 \leq 0\}$. Obviously $C_1 \bigcap C_2 = \emptyset$. It can be verified that the sequence $\{x^k\}$ generated by the sequential HLWB satisfies that $\|x^k\|_2 \to \infty$ as $k \to \infty$.

For the simultaneous HLWB, we give the following example.

Example 4.3. As in Example 4.1, we assume that $C_1 = \{(x_1, x_2) : x_1 \geq 1\}$, and $C_2 = \{(x_1, x_2) : (x_1+1)^2 + x_2^2 \leq 1\}$. It can be verified that the sequence $\{x^k\}$ generated by the simultaneous HLWB converges to $\omega_1(1,0) + \omega_2(0,0) = \omega_1(1,0)$, a fixed point of the operator $\omega_1 P_{C_1} + \omega_2 P_{C_2}$, where $\omega_1, \omega_2 > 0$ and $\omega_1 + \omega_2 = 1$.

From Example 4.3 and some other examples we have following conjecture.

Conjecture 4.1. *Assume that $\omega_i > 0$, $1 \leq i \leq m$, and $\sum_{i=1}^m \omega_i = 1$.*

- *If the operator* $\sum_{i=1}^{m}\omega_i P_{C_i}$ *has fixed points, then there exists fixed point* $\hat{x}$ *such that*
  (1) The $P_{C_i}((\hat{x}))$, $1 \le i \le m$, *are the all accumulation points of sequence* $\{x^k\}$ *generated by the sequential HLWB.*
  (2) The sequence $\{x^k\}$ *produced by the simultaneous HLWB converges to* $\hat{x}$.
- *If* $\sum_{i=1}^{m}\omega_i P_{C_i}$ *has no fixed points, then the sequences generated by either the sequential HLWB or the simultaneous HLWB is unbounded. In this case, Algorithm* 3.1 *also generates an unbounded sequence.*

5. Conclusions

In this note, based on the reformulations of the CFP and the steepest descent method, we further discussed several well-known algorithms for the CFP. In the linear case we show the connection between two projection algorithms for the CFP and the BAP respectively. It is well-known that the projection-type methods greatly depend on the efficient computations of the projection. To this end, some surrogate projection methods, in which the exact projection is replaced by some suitable approximate projections, were proposed to overcome the drawback of the exact projection methods.

In conclusion, the projection-type algorithms for the CFP or the BAP require less smoothness assumptions, so they are easily executed. The other advanced methods should be applied when better smoothness conditions of the functions are available.

Acknowledgments This work was supported by the National Natural Science Foundation of China, Grant 10571134 and was done in part while the author was visiting the University of Minnesota. The author would like to express his gratitude to Prof. Zhi-Quan Luo for his warm hospitality and for pointing out an error in the earlier version of the article.

References

[1] J.-P. Aubin, *Optima and Equilibria*, Springer-Verlag, 1993.
[2] H. Bauschke, *The approximation of fixed points of compositions of nonexpansive mappings in Hilbert space*, J. Math. Anal. Appl., 202 (1996), pp. 150–159.
[3] H. Bauschke and J. Borwein, *On projection algorithms for solving convex feasibility problems*, SIAM Rev., 38 (1996), pp. 367–426.
[4] D. Bertsekas, *Constrained Optimization and Lagrange Multiplier Methods*, Academic Press, New York, 1982.
[5] J. Bonnas, J. Gilbert, C. Lemaréchal, and C. Sagastizábal, *Numerical Optimization*, Springer- Verlag, 2003.
[6] C. Byrne, *A unified treatment of some iterative algorithms in signal processing and image reconstrution*, Inverse Probl., 20 (2004), pp. 103–120.
[7] Y. Censor, *Computational acceleration of projection algorithms for linear best approximation problem*, Linear Algebra Appl., 416 (2006), pp. 111–123.
[8] Y. Censor, M. Altschuler, and W. Powlis, *On the use of Cimmino's simultaneous projections method for computing a solution of the inverse problem in radiation therapy thertment planning*, Inverse Probl., 4 (1988), pp. 607–623.

[9] Y. Censor and T. Elfving, *Block-iterative algorithms with diagonally scaled oblique projections for the linear feasibility problem*, SIAM J. Matrix Anal. Appl., 24 (2002), pp. 40–58.

[10] Y. Censor, A. De Pierro, and M. Zaknoon, *Steered sequential projections for the inconsistent convex feasibility problem*, Nonlinear Anal., 59 (2004), pp. 385–405.

[11] Y. Censor and S. Zenios, *Parallel Optimization: Theory, Algorithms and Applications*, Oxford University Press, New York, NY, USA, 1997.

[12] P. Combettes, *Construction d'un point fixe commun à famille de contractions fermes*, C.R. Acad. Sci. Paris, Sèr. I Math., 320 (1995), pp. 1385–1390.

[13] P. Combettes, *The convex feasibility problem in image recovery*, Adv. Imag. Elect. Phy., 95 (1996), pp. 155–270.

[14] P. Combettes, *Hilbertian convex feasibility problem: Convergence of projection methods*, Appl. Math. Opt., 35 (1997), pp. 311–330.

[15] P. Combettes, *A block-iterate surrogate constraint splitting method for quadratic signal recovery*, IEEE T. Signal Process., 51 (2003), pp. 1771–1782.

[16] G. Crombez, *Finding projections onto the intersection of convex sets in Hilbert spaces*, Numer. Func. Anal. Opt., 16 (1995), pp. 637–652.

[17] F. Deutsch, *Best Approximation in Inner Product Spaces*, Springer-Verlag, New York, NY, 2001.

[18] F. Deutsch and H. Hundal, *The rate of convergence of Dykstra's cyclic projections algorithm: The polyhedral case*, Numer. Func. Anal. Opt., 15 (1994), pp. 537–565.

[19] F. Deutsch and I. Yamada, *Minimizing certain convex functions over the intersection of the fixed point sets of nonexpansive mappings*, Numer. Func. Anal. Opt., 19 (1998), pp. 33–56.

[20] J. Eckstein and D. Bertsekas, *On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators*, Math. Program., 55 (1992), pp. 293–318.

[21] U. García-Palomares, *Parallel projected aggregation methods for solving convex feasibility problem*, SIAM J. Optimiz., 3 (1993), pp. 882–900.

[22] J.-L. Guffin, Z.-Q. Luo, and Y. Ye, *Complexity analysis of an interior cutting plane method for convex feasibility problems*, SIAM J. Optimiz., 6 (1996), pp. 638–652.

[23] G. Herman, *Image Reconstruction from Projections: The Fundamentals of Computerized Tomography*, Academic Press, New York, NY, USA, 1980.

[24] H. Hundal and F. Deutsch, *Two generalizations of Dykstra's cyclic projection algorithm*, Math. Program., 77 (1997), pp. 335–355.

[25] A. Iusem and A. De Pierro, *On the convergence of Han's method for convex programming with quadratic objective*, Math. Program., 52 (1991), pp. 265–284.

[26] K. Kiwiel, *Exact penalty function in proximal budle methods for constrained convex nondifferential minimization*, Math. Program., 52 (1991), pp. 285–302.

[27] K. Kiwiel, *Block-Iterative Surrogate projection methods for convex feasibility problems*, Linear Algebra Appl., 215 (1995), pp. 225–259.

[28] K. Kiwiel, *Monotone Gram matrices and deepest surrogate inequalities in accelerated relaxation methods for convex feasibility problem*, Linear Algebra Appl., 252 (1997), pp. 27–33.

[29] K. Kiwiel and L. Bozena, *Surrgate projection method for finding fixed points of firmly nonexpansive mappings*, SIAM J. Optimiz., 7 (1997), pp. 1084–1102.

[30] Z.-Q. Luo and J. Sun, *Polynomial cutting surface algorithm for the convex feasibility problem defined by self-concordant inequalities*, Comput. Optim. Appl., 15 (2000), pp. 167–191.

[31] L. Marks, W. Sinkler, and E. Landree, *A feasible set approach to the crystallographic phase problem*, Acta Crystallogr. A, 55 (1999), pp. 601–612.

[32] L. Qi, *Superlinearly convergent approximate Newton method for LC^1 optimization problems*, Math. Program., 64 (1994), pp. 277–294.

[33] Y. Nesterov and A. Nemirovskii, *Interior-Point Polynomial Algorithms in Convex Program-*

ming, SIAM, Philadephia, 1994.

[34] Y. NESTEROV, *Smootning minimization of non-smooth functions*, Math. Program., 103 (2005), pp. 127–152.

[35] J. PEÑA AND J. RENEGAR, *Computing approximate solutions for convex conic systems of constraints*, Math. Program., 87 (2000), pp. 351–383.

[36] Q. YANG, *The relaxed CQ algorithm solving the split feasibility problem*, Inverse Probl., 20 (2004), pp. 1261–1266.

[37] Q. YANG, *On variable-step relaxed projection algorithm for variational inequalities*, J. Math. Anal. Appl., 302 (2005), pp. 166–179.

[38] J. ZHAO AND Q. YANG, *Several solution methods for the split feasibility problem*, Inverse Probl., 21 (2005), pp. 1791–1799.